

### Changes in data collection procedures for process-generated data and methodological implications: the case of ethnicity variables in 19th century Norwegian censuses

Thorvaldsen, Gunnar

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

#### Empfohlene Zitierung / Suggested Citation:

Thorvaldsen, G. (2009). Changes in data collection procedures for process-generated data and methodological implications: the case of ethnicity variables in 19th century Norwegian censuses. *Historical Social Research*, 34(3), 168-190. <https://doi.org/10.12759/hsr.34.2009.3.168-190>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more Information see: <https://creativecommons.org/licenses/by/4.0>

# Changes in Data Collection Procedures for Process-Generated Data and Methodological Implications. The Case of Ethnicity Variables in 19th Century Norwegian Censuses

Gunnar Thorvaldsen \*

**Abstract:** »Verfahren zur Harmonisierung von prozessproduzierten Variablen und methodologische Implikationen. Das Beispiel der „Ethnizität“ in norwegischen Volkszählungen des 19. Jahrhunderts«. This article discusses ethnic classification in the censuses in order to prepare its use as an independent variable in for instance demographic studies. The availability of census data and other public administrative data are increasing, also cross-nationally. In order to use these consistently in analyses, variables and categories have to remain the same over all measurement points, and the same type of person should whenever possible be classified and categorized in the same way. Using the case of ethnicity variables in Norwegian censuses, the article a) illustrates that with process-produced data, the contents of the original manuscripts are not necessarily comparable over time and space; b) it then discusses factors leading to these incompatibilities and c) suggests how to harmonize the inconsistencies.

**Keywords:** Longitudinal Analysis, Process-Generated Data, Social Bookkeeping Data, Public Administrative Data, Institutional Filters, Measurement, Census, Ethnicity.

## 1. Introduction

Ethnicity is a crucial independent variable when explaining demographic and social history differentials in the ethnically mixed parts of Fennoscandia. Computerized census manuscripts contain ethnic markers for Norway during the period 1865 to 1910, which can be used both in both cross-sectional and longitudinal studies of the ethnic groups. Since the data are process-generated, however, they are not compatible across time and space. Below I shall illustrate these incompatibilities and their origins with examples from Northern Norway and suggest methods for handling them.

Until the 1960s the southern part of Norway was ethnically homogeneous, except for the in-migration of German and Dane administrators until 1814,

---

\* Address all communications to: Gunnar Thorvaldsen, Norwegian Historical Data Centre, Faculty of Social Sciences, University of Tromsø, N 9014 Tromsø, Norway; e-mail: GunnarTh@sv.uit.no.

Finnish farmers and a Swedish proletariat during the late 19th century. In the northernmost three or four provinces, however, three ethnic groups lived together over the centuries in what has been coined the meeting place of the three tribes (Skøyen 1918): Sami, Norwegians and Fins, mentioned in the order they came to settle the region. There are clear indications of regional differentials with respect to demographic factors (Sogner 1979) and also such differences between the ethnic groups (Minde 1975). This article aims to discuss ethnic classification in the censuses in order to prepare use of this variable in for instance demographic studies. We should be aware, however, that this is a two-way process: demographic differentials will also affect how the ethnic groups are composed and defined especially through migration and inter-marriage. We should think of the analogy, therefore, that in this kind of study we are standing on a moving platform while aiming at a moving target.

The intricacy and even vagueness of the concept of ethnicity is one reason why another article about it is in its place, in spite of the many who have already written about ethnicity in the Norwegian censuses (Kiær 1882, Otnes 1979, Thorsen 1972, Torp 1986, Bjørklund 1985, Thuen 1987, Steinlien 1986, Hansen and Meyer 1991, Hansen 1994, Jåstad, 2003). Most recent publications on ethnicity in the censuses is oriented towards the local level, while the present article aims to overview how the system for assigning ethnicity in the census was designed and redesigned on the national level. Omissions in previous accounts is another reason; for instance they do not refer to the detailed rules for aggregating census results about the ethnically mixed population, or how this effected the layout of the census manuscripts from different years.

Also, it is only now that an overview of ethnicity and other variables can be based on access to computerized transcriptions of four complete, nominative censuses for the ethnically mixed areas; 1865, 1875, 1900 and 1910 (cf [nappdata.org](http://nappdata.org)). The existence of these sets of microdata, which are distributed for research internationally, makes it mandatory to provide a thorough description of the ethnicity variables contained in the microdata in a major language, so that researchers using them can understand the background and methodology for their collection. There are indications of local differences in how ethnicity was ascribed to population segments in the censuses over time, making further locality studies of these processes desirable. Future community studies should profit from an updated overview on the national level.

Few researchers any more believe in straightforward definitions of ethnicity which categorize people by singular criteria such as physical characteristics. Left behind are the melting pot kind of definitions where multiple ethnic groups as the result of migration become one homogeneous flock, in the most extreme versions justified as necessary for successful nation building (Turner 1893/1966). Neither do we any more see ethnic heterogeneity as a continuous struggle between such groups or the alienation of the in-migrants or the small-numbered groups (Handlin 1951). Favoured today are definitions analogous to

the salad bowl, with ethnicity as something dynamic and multifaceted, resulting from the bringing together of different cultures and gene pools.

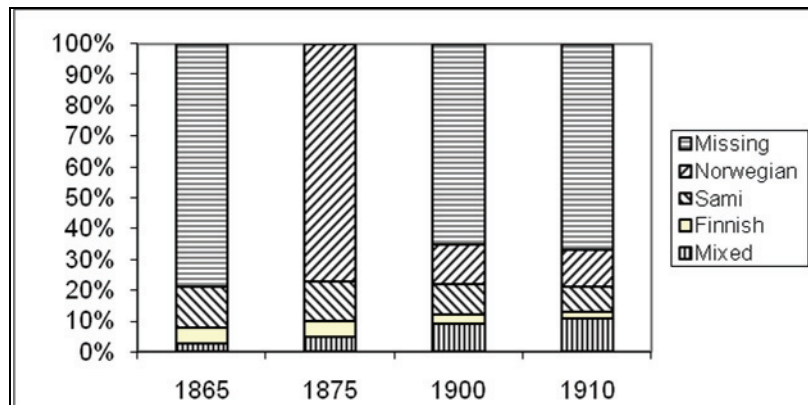
It is indeed a signal about the complexity of the issue that the word ethnicity was not used in the census forms, instructions or publications of aggregates during the period covered here. Instead this population characteristic was called nationality or "Nationalitet" in the vernacular. During a century when Norway was in a flux of national consolidation from being a part of the Danish Kingdom via confederacy with Sweden to independent sovereignty, this wording was hardly a coincidence (Patriarca 1994). The integration of subgroups such as national minorities was a priority in many European countries, and the "enthusiasm for social counting" was part of this process already from the 1830s (Westergaard 1932, 136ff). Later in the century the international statistical conferences adopted several recommendations on how to classify national minorities in order to make the statistical aggregates multilaterally comparable (Lie 2001, 147). This concept of nationality used in the Norwegian 19th century censuses is more in line with a German nationalism based on consanguinity than with the concept of the nation based on a territorial and political unit springing out of the French revolution (Hobsbawn 1990). Internationally there was a development, for instance in Britain to reserve the nationality concept for territorial nations rather than co-ethnic groupings (Manning 2005, 141f). The Central Statistical Bureau was quite conservative in its use of the nationality concept, and this may have affected the compilation of the ethnicity variables. Only as late as 1930, in the last census to note ethnicity on the individual level, the word race was used instead, presumably with fewer connotations at the time than in later decades.

A further problem is that it is usual for independent variables to only explain small or marginal effects in demographic studies. What does for instance a difference of a couple of percentage points in infant mortality between ethnic groups mean if ethnicity is murkily defined and perhaps also confounded by other variables such as place of settlement? (Jåstad 2003). A further problem is how the effect of ethnicity can be compared in a meta-analysis of several previous projects, if the concept is defined differently in different studies?

The article throws light on these questions with examples from late 19th century and early 20th century censuses taken in Northern Norway. The censuses 1865 through 1910 make up the only computerized microdata available from north-western Europe which in principle provides ethnic markers for the whole population. In addition there are aggregate statistics for the period 1845 to 1930. This article will summarize how the principles behind ethnic classification developed from census to census as expressed in the instructions to census takers and the categorization in the ethnic aggregates. Next it discusses how the problems which necessarily are inherent in ethnic classification were dealt with. While most previous analyses have limited themselves to analysing ethnic markers in a locality such as a parish, this article intends to present the

national system for ethnic classification. For an example of ethnic distribution in the original censuses, cf figure 1.

Figure 1: Ethnic distribution in census manuscripts for Troms province.



## 2. Source Material

Norwegian regional and national archives hold nominative sources covering the country's population during more than three centuries. While the ministerial records provide scanty evidence about ethnicity, there is little or no such information in the oldest male, nominative or statistical censuses taken from the 1660s until 1835. During the century-long period from 1845, however, administrators aimed to map systematically the ethnic characteristics of each individual, household or larger group. This paper shall mainly concentrate on the nominative and computerized censuses from 1865, 1875, 1900 and 1910 since the earlier ones are statistical only and the later ones are less readily available for analysis due to privacy restrictions in the Laws of Statistics from 1907 and 1988. The complete and encoded 1865 and 1900 censuses and the representative 1875 census are available together with constructed variables through the North Atlantic Population Project (NAPP). The census from 1910 has been transcribed for Norway but is so far only partly encoded for statistical analysis and may not be made generally available until December 2010.

Geographically I shall focus on the provinces of Nordland, Troms, and Finnmark, Norway's three northernmost provinces which lie within the ethnically mixed part of the country where the extra questions about ethnicity were asked. We should in addition remember the Sami groups in the mid-Norwegian Trøndelag provinces and the Fins in the south-eastern province of Hedmark. Since the chief goal of any census is more cross-sectional than chronological, their variables were defined more to be comparable within the nation and inter-

nationally than to render how the characteristics of the population changed over the decades. This is especially true for the ethnicity variables, which may be more ephemeral and difficult to define than the contents of any other field in the census questionnaires. Thus, the ethnic categories were constructed both with a view to suit political and administrative needs in any particular census year and to improve the ability of the census to mirror the development of the nation's ethnic composition, while often sacrificing comparability with previous censuses.

Thuen (1987) has primarily studied how the three specific ethnic categories were replaced by the mixed categories or the under-reporting of ethnicity in late 19th and especially early 20th censuses for specific localities north of Tromsø. Hansen and Meyer (1991) have compared the late 19th century ethnic markers in censuses with those for the same individuals in the ministerial records for a parish south of Tromsø, finding that many who were denoted "mixed" in the ancestry based census were classified by the priest as Sami in the ministerial records on cultural or language grounds. Torp (1986) was more critical in an early work, but has later found evidence that the lack of consistency between the ethnicity marked in different censuses can indeed be explained with further empirical evidence. E.g. the ethnic "chameleon" who was Sami and Fin in different censuses turned out to be of Finnish stock, but had immigrated to Norway with the Sami (Torp 1990). He is thus in line with Barth's work (1969) on the shifting ethnicity of individuals who crossed one or more ethnic borders during their lifetime. Several single individuals have been identified who for instance were denoted as Sami while living in a Sami community, but with a different ethnic marker after having moved to a predominantly Finnish or Norwegian area. This may be caused by the census takers noting ethnicity on the group rather than the individual level, but it can also be caused by the migrants changing their clothe-style and their main language after moving (page 186: "the in-migrating Norwegians after some time adapt to the Sami language, clothing and life style, caused partly by the custom that the domestic work is done according to the wife's ethnicity" (Thorsen 1865). A general conclusion is that the census' ethnic marker is of better quality than according to hearsay (Hansen and Meyer 1991, 50, Torp 1990).

Research into the classification of ethnic groups must be informed about the more or less open conflicts between them. There are many reports about mutual attacks on animals, particularly dogs, and even a few instances of manslaughter. (Niemi 2000, Haugli 1981, Thorvaldsen 2004). The most well-known conflict is the Kautokeino uprising in 1852, when the bailiff and the merchant were killed and the parson flogged by the insurgents. It is noteworthy that the former and the latter held offices directly responsible for census-taking in their district and that the uproar happened at the time when ethnicity was a new variable in the census. Kautokeino is a predominantly Sami municipality on the mountain plateau in the middle of Finnmark province. Two of the insurgents were de-

capitated and fifteen others were sentenced to long penal servitude, several of whom died in jail. Less well-known is the conflict over the immigration of Finns towards the Finnmark coast during the severe starvation crisis in Finland in 1868. A group of Norwegians published a letter in a Finnmark newspaper urging the authorities to dispatch extra troops to protect the local food supply against potentially marauding, hungry Finns. The authorities fortunately restricted themselves to publishing a warning in some Swedish and Finnish newspapers that new Finnish immigrants risked being sent back to Finland, which was sometimes done. There was obviously no need for troops, and other Norwegians answered with a campaign to send emergency aid to Finland (Myhre 2003, 203).

These changing aspects of ethnicity makes it a most complicated census variable, it is thus typical that the publication series *Historical Statistics* from Statistics Norway (1994) with seven editions contain no results about ethnicity – data collection on this category is hardly mentioned in the preface, and it was considered too complicated to construct long-term aggregate time series from this variable (cf <http://ssb.no/histstat/>). The present article will attempt to model the ethnicity concept used in the Norwegian censuses along three dimensions. Ethnicity can be defined

- *on the individual and group levels*: ethnicity can be stated for each individual in the census or for a group such as a family.
- *by subjective versus objective criteria*: is the ethnicity marker decided by the individuals themselves or by the census takers?
- *on the basis of language and other cultural characteristics as distinguished from the persons' ancestry*: the census taker could note ethnicity based on the language spoken or dress worn, or alternatively be imputed from the ethnicity of the parents.

This has a direct bearing on the contents of the census manuscripts and the contemporarily published or newly reconstructed aggregates also because the methodology recommended in the instructions provided to census takers changed over time (cf <http://www.rhd.uit.no/nhdc/census.html>). The statistical nature of the pre-1865 material would favour a group based criterion since in principle no individual information was supposed to be provided or preserved. Statistics on ethnicity was created for the first time in the 1845 census. Unlike gypsies who had no permanent residence, the Sami and the Finns were counted as part of Norway's population.

Ethnicity aggregates were given extra significance by their inclusion in the procedural introduction to the census results. Less than ten percent of the 14464 Sami lived in the southernmost province of Nordland, an obvious undercount since many were on their winter pastures in Sweden when the census was taken in early January. While few Sami lived in Southern Norway, nearly one third of the 4425 Finns were enumerated here; in Hedmark province where several Finnish settlers had cleared farms in the vast forests not far from the

border with Sweden. Next to no Sami were found in towns while 378 Finns were enumerated in urban places. This is again a virtual undercount, since there were whole colonies of Finns in urban and suburban-like settlements, especially inside and outside Vadsø (Niemi 1977). As a result of the 1855 census, a table with the number of Sami (nomadic and resident) and Finns in each town or parish was published. Since the instructions to census takers are so limited, we shall also compare the more comprehensive rules behind the published aggregate tables over time.

### 3. Ethnicity in the Nominative Censuses 1865 to 1930

The census authorities repeatedly changed the questionnaires, the instructions and the practice of census taking every time they planned a new census. This was done in order to make the new census more accurate than the previous, in order to follow international recommendations and to provide information needed by the authorities. Both for the late 19th and early 20th century we shall overview how when measuring the ethnicity variable, censuses varied along several dimensions:

- 1) Collection of Data on (Head of) Housholds vs. on Individuals
- 2) The Respondents' Native Language
- 3) Data Collection by Census Takers vs. Self-Enumeration
- 4) Under-Enumeration
- 5) The Wording of the Variable: Nationality – Ethnicity – Race
- 6) Classification on Language vs. Ancestry

#### 3.1 Head of Houshold vs. Individual

The ethnic classification for a whole group might be based on one significant member of the household or farm community, for instance the (usually male) head of household. This may apply also to the nominative 1865 census, since here ethnicity was reported, not in a special column, but by splitting families or households by group according to the ethnicity of each person. It made things easier for the census taker if an ethnically mixed family was not split but rather noted together in the manuscript and assigned to the ethnicity of the main person. Since Norwegians should be noted first, followed by Finns and finally by the Sami, we may suppose that especially information about the latter group was underrepresented, and that it was not always noted if they were nomadic. Census takers were instructed to also note children's mixed ethnic origin, but since this would multiply the number of groups, it was not always done. After the information in the 1865 census manuscripts was transcribed and computerized, ethnically mixed families and households were reconstructed manually, and the information about ethnicity was copied from the group headings to each person. On the basis of the resulting edited census manuscript each indi-



vidual's ethnicity has been encoded in order to be more comparable with the later censuses. Copying information from the group to the individual level was thus our first action to harmonize these process-generated data sets.

### 3.2 The Respondent's Native Language

In addition to ethnicity, census takers already in 1865 were told to note language in the remarks column, more specifically whether the person understood Norwegian. It is unclear, however, if this applied only to ethnically mixed persons and also what level of language skill was required. Since the language comments have been encoded together with the rest of the microdata, we can easily see that census takers interpreted the instructions quite verbatim. Nearly all the comments refer to the person's command of Norwegian, telling whether he or she could speak or understand Norwegian, sometimes also in combination with other languages. Three times more persons were noted as having no command of Norwegian than were noted as speaking Sami or Finnish – which must have been the case for those who did not know Norwegian. The majority of the Sami and Finnish got no comment about language and the clear majority of the rest were put as understanding Norwegian. It is unclear from the 1865 manuscripts whether this means that most of them did not have a command of Norwegian or the more likely interpretation that the census takers forgot to note their language skills. In the tabulation reports from later censuses, the census authorities commented that the 1865 language variable was never analysed statistically, which is only natural given that it is incompatible with how language was reported in other censuses.

The same authorities seem to have forgotten however, that language was also reported in the 1875 census manuscripts (Torp 1986, 71). The instructions told census takers to provide in the comments field for “all Finns and Sami and for persons of mixed ethnicity information about which *language they usually spoke*”. Such notes were made for some persons in the northern parts of Nordland and more consistently in Troms and Finnmark provinces. With 3674 persons of mixed ancestral ethnicity only in Finnmark the extra information about their preferred daily language could obviously be helpful when specifying ethnicity, but the instructions leave us in the dark with respect to if or how language was actually used in 1875. This year the census tabulators could base their work on the detailed reporting of ancestral ethnicity for both parents. The number of persons counted as mixed more than doubled from 1855 to 1875, and much of this increase happened after 1865. Much of the increase was due to the more detailed and individual oriented way of noting ethnicity in the censuses over time. From 1875 onwards, ethnicity was noted in a special field on the census form, which should inspire the census taker to ask about each person's individual ethnicity in particular. The inclination to do so may have been especially marked in 1891 when one separate form was to be filled for

each individual – a cumbersome format which explains why this census has not yet been computerized.

### 3.3 Census Takers vs. Self-Administration

Starting with 1865, census taking in towns shifted to self-enumeration while the system of hired census takers, often teachers, was continued in rural parishes and municipalities. The parish was the basic administrative census unit and the parson was in charge of the census until 1865. The bailiffs were the main administrators in *municipalities* after 1875. Self-enumeration must not be interpreted literally in a society where the teaching of writing skills in schools had been introduced relatively recently. For one town it has been shown how a few men could fill in the forms for whole blocks of buildings (Drake 1991), while in other instances the job was done by the head of household or the property owner.

In theory the introduction of self-enumeration should shift ethnic classification in the direction of subjective criteria. By this is not intended the subjective opinions of the census takers, but rather the opportunity for people to state their opinions about their own ethnicity. The effect was in actual practice rather limited, though. Not only did some persons fill in many forms, the urban population was rather homogeneous. It has for instance been reported how immigrants from Finland upon arriving in Tromsø where sent by the first boat available eastwards to the village of Skibotn – with Finland within walking distance.

### 3.4 Under-Enumeration

Still, the generally subjective element inherent in self-reporting could influence the choice of ethnic markers, since the persons were supposed to be present and asked by urban and rural census takers about the various items in the census questionnaire. We may assume that children were asked less frequently and guess that women were less likely to give their opinion. Also, the circa three percent of the population who were absent on census day could not be interviewed, but from 1875 they should be asked about their ethnicity if temporarily present in the ethnically mixed part of Norway. This ought to have reduced the undercount among nomadic Sami, most likely also among the many Finns who were geographically mobile. Comparing how ethnicity and other characteristics were noted differently for the same person at the same time in two different places is a task that awaits future research (Thorvaldsen 2006).

### 3.5 The Wording of the Variable: Nationality – Ethnicity – Race

As mentioned in the introduction, the concept of ethnicity was not introduced during this period. Instead the word “Nationalitet” was used both in the instructions, the questionnaires and the printed aggregates. Thus, “the Sami Nation”

could be referred to (Otnes 1979). From 1875 a new column was provided for citizenship, appropriately headed “Hvilken Stats Undersaat” – not to be filled for citizens of Norway. We should be open to the possibility that some Sami meant their nationality was Norwegian (like we understand this concept today), and wanted this to be expressed in the “Nationalitet” column – cf the low number of Sami enumerated in the towns (Aubert 1978).

### 3.6 Classification based on Ancestry vs. Language

In the 1875 census wards where the population besides Norwegians consisted of a significant part of Finnish or Sami people, information on the inhabitants' ethnicity should be reported in the questionnaire's 3rd field (after the person's name) according to the formula N=Norwegian, L=Sami, K=Finnish, B=Mixed. While the self-reporting subjective element may apply only infrequently and the individual versus group problem can be adjusted for with access to person level microdata, shifts between definitions of ethnicity based on ancestry and culture will be more difficult to handle in the Norwegian census material. According to census instructions the individual ancestry should be used in all the nominative censuses. This is especially clear in 1875, when parents' ethnicity was marked with a two-letter code, the first letter indicating the ethnicity of the father, the second giving that of the mother. Thus, when children were living with their parents we have data for three generations, and in case any grandparents were present a four generation ethnic lineage is indicated. This interesting combination provides good opportunities for checking internal consistency of reporting ethnicity within each family.

Persons whose father or mother had mixed ethnicity are problematic, however, since these will be denoted by the letter B, which cannot be qualified unless a parent or grandparent is also present. This problem was less severe with the more unified reporting of ethnicity in 1865. The census takers could make their job easier also in 1875 by entering the information as if the household was homogeneous, not only could the ethnicity marker be given for the whole group, they did not have to enter information on language in homogeneous households. Perhaps the organizers of the census had this in mind when they specifically ordered the bailiffs to control whether the information on ethnicity had been entered before making their local aggregates. The grouping together of people with the same ethnicity would inspire the census taker to note the primary ethnicity either because a majority of the ancestors belonged to one specific group, because of patriarchal criteria or because cultural factors played into the choice of ethnic marker

The cultural element increased its importance in the two last censuses of the 19th century, but it is unclear to what degree. A special field for language was included from the 1891 census onwards, but language was quite often reported both in 1875 and 1885 and even sometimes in 1865, then as marginal com-

ments. Language reporting entails its own problems; according to instructions the language ordinarily used in the home should be noted. In 1865 the instructions focussed on knowledge of Norwegian, and we do not know how often the census taker also later noted that the persons spoke Norwegian since this often was the language used to communicate with him. The 1891 instructions tried to amend this problem somewhat by defining language as the one spoken in the person's home, and this should only be noted if different from the person's ethnicity. However, there is no special instruction on how to define ethnicity in 1891, the census taker should simply underline one of the three main groups or "mixed" and the Sami should be further classified as nomadic or permanent residents. The census authorities may have hoped that more systematic gathering of information on language should solve the problem of how to deal with the growing groups of ethnically mixed persons due to mixed marriages.

When tabulating the information for the aggregate publications, the ethnicity issue came out as more complicated, as can be seen from an appendix to the tables about the mixed groups listing ten rules for how to deal with them.<sup>1</sup> The rules can be summarized as follows: a) Where ancestral ethnicity was consistent, this should be used regardless of language spoken. b) Persons of mixed ethnicity could be classified by language, especially if their parents' ethnicity was not known. c) If persons of mixed origin could speak both Norwegian and another language, they should be tabulated as Norwegian. d) Multilingual persons of mixed Sami-Finnish ancestry should be classified according to their fathers' ethnicity. If father's ethnicity was unknown, persons who were of mixed Sami-Finnish origin and spoke these two languages should be classified according to the dominating ethnicity in their household. Thus, the rules still kept ancestral ethnicity as the main criterion since language was only to be

---

<sup>1</sup> Rules for distributing people by ethnicity in 1890, 1900 and 1910 census aggregates

As Norwegian are classified in addition to the pure Norwegians:

1. Mixed of Norwegian-Sami and Norwegian-Finnish and Norwegian-mixed, who spoke Norwegian, Norwegian and Sami, Norwegian and Finnish, as well as Norwegian, Sami and Finnish.
2. The mixed, whose parents' ethnicity was missing and who spoke Norwegian, Norwegian and Sami, Norwegian and Finnish, as well as Norwegian, Sami and Finnish.

As Finnish (Quaines) are classified:

1. Pure Fins and mixed of Finnish-mixed regardless of language.
2. The mixed of Sami-Fin, whose language was Finnish, Norwegian and Finnish or Norwegian, Sami and Finnish, but in the latter case only if the father's ethnicity was Finnish.
3. The mixed of Norwegian-Fin, whose language was only Finnish, as well as
4. The mixed whose parents' ethnicity was missing and who spoke only Finnish.

As Sami are classified:

1. Pure Sami and the mixed of Sami-mixed regardless of language,
2. The mixed of Sami-Fin, whose language was Sami, Norwegian and Sami, or Norwegian, Sami and Finnish, but in the latter case only if the father's ethnicity was Sami,
3. The mixed of Norwegian-Sami, whose language was only Sami, as well as
4. The mixed, whose parents' ethnicity was missing and who spoke only Sami.

used for the ethnically mixed and when ethnic markers were missing. Also, one parent's pure ethnicity could be decisive for classifying ethnicity, regardless of language.

The classification was biased towards using Norwegian ethnicity if one parent was Norwegian and the person spoke some Norwegian. Language remained a subsidiary criterion for unclear cases, and could also be overruled by the group criterion. In one table it was also attempted to distribute ethnically mixed persons according to what fraction of their ancestors belonged to the respective groups. Thus a person with a Norwegian father and both Sami and Finnish grandparents on the mother's side would count as 1/2 Norwegian, 1/4 Sami and 1/4 Finnish. Since in 1891 the variables of (ancestral) ethnicity and primary language were – in principle – enumerated separately, they could be cross-tabulated in the aggregate publications. In addition, ethnicity was analysed by municipality of residence, age, marital status and finally by occupations for the Sami and the Finns separately.

The direct background for the attempt to use the language criterion more systematically was the influence of the international statistical conferences whose importance grew and where Norwegian representatives participated eagerly through the last quarter of the nineteenth century. As one of their efforts to standardize census questions and variables for international, cross-sectional comparison, they recommended census-taking nations to base the classification of ethnicity on cultural traits such as language, rather than ancestry. This change fitted well into contemporary official policies with regard to the ethnic minorities: the Sami and the Fins should be incorporated into the Norwegian nation as quickly and completely as possible. They should for instance not be allowed to speak their vernacular while attending school. The use of ancestry when reporting the size of the ethnic groups in the census could hardly map the progress of the wanted development, but reporting based on cultural criteria could.

There was also a more practical reason, since as the mixed population element grew it became increasingly difficult to classify ethnicity by using only ancestry. People might not know the exact ethnicity of their grandparents, and the number of different mixed subgroups became unmanageable. Fortunately, the reform came too late and too slowly to replace the information based on ancestry. In the census manuscripts from the turn of the century we can study how the ethnic markers were changed on the basis of the language variable when they were prepared for aggregation in the Statistical Bureau. The crossed-out, original marker is usually legible, and has been transferred to the computerized microdata.

### 3.7 How this Reflects in Statistics

The aggregates based on the two first nominative censuses reported an increase in the Sami population of ten percent from 1855 to 1865 and by six percent during the next decade. The fact that this is the same proportional growth as the rest of the population might increase our trust in the data as was commented by The Table Office upon publishing the 1875 results. We should show greater caution, however, since the decline in growth for the second decade was due to emigration. Since we know the Sami had low emigration rates, we might rather worry that they were underenumerated. We do not know how many persons were not marked as Sami or Finnish in 1875 because they lived in municipalities where the questionnaires without an ethnicity column were used, but the census authorities saw it as a problem and from 1891 provided some ethnicity forms in all municipalities.

The resulting aggregates list Finns in six and Sami in twelve of Norway's nineteen provinces, but their numbers outside the core areas were small – five Finns and 446 Sami still indicate that some underenumeration was likely (NOS III, no 278, tabell 1). It has been maintained that there was a real reduction in nomadic reindeer-herding Sami on the mountain plateaus due to a delayed effect of the closing of the borders around 1850. Some could not trek to Norway any longer while others risked losing their Sami identity in the process of giving up their reindeer. The increase in the Finnish group was 36 and 28 percent respectively for the two inter-census decades. On the basis of the birth place information in the 1865 and 1875 censuses it was calculated that twelve percent of the increased Finnish group was due to immigration. Since this included birthplaces in Finland only and many ethnic Fins came from north-eastern Sweden, the relative effect of immigration ought to be increased somewhat (Kjær 1882, 149f). Also, many Finns went on to America via Norway, some after they had been noted in a census.

The aggregate results on ethnicity were more specific in 1875 than ten years earlier, but still did not tabulate any language or occupation criterion for the Sami or Finns. Instead they stressed geography, giving the number of Sami, Finns and mixed (in the French language aggregates called “Métis”) by sex on the level of the municipality. In addition to Northern Norway and North Trøndelag results were included for seven municipalities in South Trøndelag and for three municipalities settled by Finns in Hedmark province. In the latter places no Sami were enumerated, as expected, but the tabulation of only 20 – twenty – Sami persons in the towns in Northern Norway should make us think seriously about under-reporting of this ethnic group. The priority given to ancestry as ethnic marker in this census was accentuated by the extra table giving the *father's* ethnicity for men and women of mixed ethnicity by province (Statistics Norway 1878-81).

When the national census scheduled for 1885 was moved to 1891 in order to synchronize census-taking with other nations, it was decided to organize a

census in the fast-developing towns. The 1885 questionnaire is quite equal to the one used in 1875 as far as ethnicity is concerned, there are two extra columns in the forms designed for the ethnically mixed north, one for the ethnicity of the father, the other for the mother. It is interesting that the 1885 census in the eastern part of Finnmark province was extended to cover also the rural municipalities. This was allegedly done for national security reasons, in order to control the thousands of Finns who had immigrated from Finland, at the time part of the Russian Empire (Eriksen and Niemi 1981, Lie 2001). These results were not made public, however, and might explain why the instructions on how to fill in information on ethnicity was scarce (“fill in Norwegian, Finnish, Sami or Mixed”). After all, most non-Norwegian ethnic groups lived outside the towns. Transcribed 1885 manuscripts are available online for Tromsø, Tana, Polmak and Nesseby municipalities, including ethnic and some language markers.

## 4. The Early 20th Century Censuses

Due to the influences from the international statistical conferences and domestic needs for more specific aggregates, census procedures and instructions were modified and standardized further during the 20th century. Still, some elements from the 19th century census taking survived in the early 20th century nominative manuscripts. Below we shall see how the ethnicity variable varied along the following dimensions during this period according to:

- 1) Reporting and Documentation Practice
- 2) New Ethnic Categories
- 3) The Respondent’s Native Language
- 4) The Wording of the Variable: Nationality – Ethnicity – Race
- 5) Classification on Language vs. Ancestry
- 6) Changing Borders.

### 4.1 Reporting and Documentation Practice

Little was changed in the census takers’ instructions and questionnaires between the 1900 and the 1910 censuses. There were the same two columns for ethnicity and language, with the same alternative categories and abbreviations. However, according to the 1910 census report, the tabulating procedure was modified somewhat, most likely to save space on the Hollerith punch cards which had been introduced when tabulating the 1900 results. Instead of electric counting, the information about returnee emigrants, ethnicity, language and some other variables only used for subgroups in the population, was copied onto handwritten individual forms which were sorted and counted manually. This is analogous to the individual level questionnaires used for all information in the 1891 and 1920 censuses. In 1910 the editing could be done on the auxil-

iary forms after the information had been copied from the original manuscripts, and the former have been discarded together with the punch cards. This explains why we find so few instances of modified ethnicity markers in the 1910 manuscripts compared with 1900.

Since we do not have access to the edited markers neither in the manuscripts nor in the transcribed version, we cannot analyse the editing process like we did in the 1900 census, for instance to find out if the use of individual cards in 1910 made the operation less oriented towards the group criterion than the tabulation from household forms in 1900. It is still clear that the tabulation of results pertaining to ethnicity and language in 1910 was meant to be as compatible as possible with the 1900 edition. The ten rules for editing the ethnic markers before tabulating them were again copied verbatim from 1891, except that the rule to copy the household's dominant ethnicity to any undefined persons was dropped.

#### 4.2 1920: New Mixed Ethnicity Categories

The procedures behind the 1920 census were less conservative both in general and with respect to the ethnicity and language questions. Lack of money for technical equipment brought back the individual questionnaire sheets used in 1891, and again the census taker could underline one of the alternative ethnic groups. In addition to Norwegian, Sami resident, Sami nomadic and Finnish three new combined options were introduced instead of the mixed category: Norwegian-Sami, Norwegian-Finnish and Sami-Finnish. The predominant ethnic component should be underlined twice. For the first time a (somewhat circular) definition of mixed ethnicity was introduced in the instructions to census takers: "Norwegians with partly Sami or Finnish ancestry should be marked as Norwegian unless at least one of the grandparents was an unmixed Finn or Sami."

#### 4.3 The Respondent's Native Language

The additional language question remained unchanged. The aggregate tabulations introduced further changes by discussing the enumeration of ethnicity in a lengthy introduction to the tables. While the 1920 instructions still used the word "Lap" for Sami, the 1920 table headings consistently called them "Finns".

#### 4.4 The Wording of the Variable: Nationality – Ethnicity – Race

Also, the word "Nasjonalitet" for ethnicity was used intermittently with "rase", the Norwegian word for race. The latter word was used systematically also in the instructions to the census takers in 1930, from which year the ethnonym for Sami was changed to "Samer" or "Sam".



#### 4.5 Classification on Language v.s Ancestry

More important, the rules for assigning ethnicity in the aggregates, which had been used in 1890, 1900 and 1910 were revised. This might superficially be seen as a most needed revision in order to simplify a set of rules too difficult to understand and too time-consuming to apply in tens of thousands of cases. However, the census authorities regarded the reform as a return to the more pure ancestry criteria applied from 1845 to 1875. They thought the 1920 and 1930 statistics were more comparable with the 1875 results than later aggregates. A number of problems with the language criterion were listed, and less weight was attached to it since only the language spoken daily should be considered and language should only be used to assign ethnicity for those who were half-mixed ("halvblandet"). If one of the three ethnic components predominated among a person's ancestors, this ancestral criterion should be decisive regardless of language. The group-criterion used in 1891 and 1900 was not reintroduced, and the rules to use language for those whose ancestral ethnicity could not be decided were dropped entirely.

The effect of the reforms on statistical results from the censuses should not be exaggerated, however. The rules about the noting of ethnicity provided for the census takers were scarce indeed. They would have to interpret them and apply them to several thousand households. There are lots of examples showing that the group criterion was applied more or less consciously, especially to household members not belonging to the family. It must also have facilitated the census takers' work if they could assume that a family or individual belonged to the ethnic group whose language they spoke, without going into details about the more or less murky ethnicity of ancestors who were no longer present, regardless of rules laid down by census administrators. The shifting between enumeration on individual sheets in 1920 and household forms in 1930 was modified by the consistent use of person cards where information on ethnicity was copied before cross-tabulating the information on ethnicity and a few other variables. Due to lack of money, fewer tables about ethnicity were published from the 1930 census than from the previous censuses.<sup>2</sup>

#### 4.6 Changing Borders

Using aggregate data to study the demography of ethnic groups over time must take a further problem into consideration: In most nations the borders of parishes and municipalities were changed more or less frequently between the censuses (Thorvaldsen 1997). Attempts to ameliorate this with correction factors based on the size of the population groups affected by the border changes

---

<sup>2</sup>. The draft tables still exist and there are plan to publish them on the Internet together with the printed material already available there, cf [www.ssb.no](http://www.ssb.no).

are problematic, since different ethnic groups tend to be located in different parts of the changed administrative area (NSD). Since ethnic minorities often live less centrally than the majority population, aggregates about them may be more affected by administrative border changes than corresponding data about other groups. Microdata are also affected by border changes if they are organized along the lines of archival principles for census manuscripts. However, microdata may be reorganized so that a consistent system of borders apply throughout the period studied. This has been done for the province of Troms by virtually “moving” farms and other settlements in the 1875 and 1900 censuses according to the administrative definitions used for the 1865 census (Thorvaldsen 1995). For the municipality of Lyngen to the east of Tromsø an even better method has been attempted whereby ethnicity has been assigned to the dominating ethnic group on each farm.

## 5. Ethnonyms and Language Problems

Both in the late nineteenth and early twentieth century confusion about ethnonyms and lack of language skills among the census takers added to the problems treated above. During the period covered in this article no census instructions or questionnaires were printed in Sami or Finnish. Neither do we have any guarantee that the census takers in the ethnically mixed municipalities could communicate with the population in any other language than Norwegian. Quite the contrary, since Sami and Finnish children were expected to speak Norwegian in school, there must have been teachers – the typical census takers – who were monolingual. For instance the directive from The Statistical Central Bureau about the 1920 census required census takers to have several qualifications, but nothing was said about language skills. Some of the more experienced teachers had learned Sami or Finnish and some had such skills due to their own family background. Other census takers may have been dependent on translators, which should be easy to find according to the high frequency of people who according to census information could speak Norwegian and Sami as well as Finnish. We have no information about the accuracy of the translation and no funds were set aside for such services according to the budgets printed in the volumes with aggregate census results. The census directives only mention higher transport costs in the North due to long distances.

In addition, a specific language problem may have distorted the census results because of how the ethnic groups were called in Norwegian. Until well into the 20th century, the Sami ethnonyms were “lapper” or “finner” (“lapons” in the French version of table titles), while the Finns all the time were called “kvæner” (“Quaines” or “Finnois” in French). The proximity between one usual Norwegian word for Sami and the more international word for Finns may have caused confusion among some census takers, since in 1845, 1855 and 1865 and in 1885 the census instructions used “Lapp” and “Finn” as synonyms

for *Sami*. In 1875, 1891, 1900, and 1910, however, the census authorities added to the confusion by using “Finsk” and “Kvænsk” as vernacular synonyms for *Finnish* in the census instructions and questionnaires. Even worse was the inconsistency in the published aggregates, where people from Finland were called “Finner” in 1890 and 1900, while the same variable name was used for Sami people in 1910 and 1920. This confusion could also affect the marking of Sami or Finnish language in the census manuscripts. Only in 1930 was a consistent terminology introduced with the modern ethnonyms “Samer” and “Kvener”.

Table 1: Ethnonyms and variable names in the census questionnaires and published tables 1865 to 1930.

Census	Field name	Sami Ethnonym	Finnish Ethnonym	Published Variable	Published ethnonym
1865	Herkomst	Fin/Lap	Qvæn	Nationalitet	Lappisk
1875	Nationalitet	Lapper	Kvæner/ Finner	Nationalitet	Kvænsk
1885	Nationalitet	Lap/Finn	Kvæn	Nationalitet	
1890	Nationalitet	Lappisk	Finsk (kvænsk)	Nationalitet	Finner $\diamond$ lapper
1900	Nationalitet	Lappisk	Finsk (kvænsk)	Nationalitet	Finner $\diamond$ lapper
1910	Nationalitet	Lappisk	Finsk (kvænsk)	Nationalitet	Finner = lapper
1920	Nationalitet	Lapper	Kvener (finner)	Nationalitet	Finner = lapper
1930	Avstamning/ Rase	Samer (finner)	Kvener	Rase	Samer = finner = lapper

A test has been programmed in order to check for this potential inconsistency by cross-tabulating the original ethnicity markers in the census with municipality and tract numbers. In the countryside, each census tract should be done by one enumerator. In the resulting cross-table two types of problems can be spotted. First whether each census taker was consistent or whether he used both “fin”, “lap” and “kven” markers in the same census (Type I problems). It will also be possible to spot tracts where the marker “fin” was used in contradiction with the census instruction (Type II problems). For instance, in 1900 “fin” should mean Finnish. If in the cross-table we can spot cells for the same tract with several persons noted as “fin”, noone noted as “lap” and several noted as “kven”, it indicates that the census taker in question used “fin” to mark Sami people, even if instructed otherwise.

The test will not work for census tracts in towns, since urban places were enumerated by several house owners, and so the test would be disturbed by any

disagreement between them. Most Sami persons lived in rural places, though. A bigger problem may be that the large variety of census markers in the 1865 census, where it is often noted for how long a person had stayed in Norway, which makes it difficult to overview the number of inconsistencies.

The test has been run for the three provinces in Northern Norway in 1875, 1900 and 1910, for the rural municipalities Tana, Polmak and Nesseby in 1885 and for Hattfjelldal, Vardø town and Båtsfjord in 1891. As expected some inconsistencies were found in the urban tracts, where both “lap”, “fin” and “kven” have been used in the same tract. Spot checks indicate that this is because some house owners used “fin” to mean Sami while others meant “Finnish” within the same tract and census. In towns, therefore, ethnicity markers should be used with special care. In the rural tracts, however, the number of inconsistencies was small. The census takers employed here were as rule educated people and the results indicate that they were well aware of the problem with the word “fin”, and possibly smiled when they noticed how the Statistical Bureau changed their definition from census to census. When they use the word “Fin”, it is often followed by “lap” or “kven” in parenthesis. In a few rural census tracts “fin” is used together with both “lap” and “kven” (type I problem), but then only a handful of persons were actually called “fin”. I have not spotted any census tract with type II problems, e.g using the word “fin” consistently in contradiction to the census instruction.

While this may indicate that the “fin” problem mostly caused confusion in urban places with small Sami openly groups, the test also revealed a third and more serious type of problem: census takers might have problems distinguishing between the Sami and the Finnish (Type III problems). In a census tract in Skjervøy municipality north of Tromsø in 1875, the ethnicity was changed from Finnish to Sami for 197 persons out of 663. This often pertained to all the persons in a family, but sometimes only some persons in a household were edited. While the error was corrected in this census tract, it should alert us to similar errors being made in other tracts, where they might not have been discovered and corrected. The type III problem cannot be detected with my statistical test unless spotted by the census taker or his administrators at the time of the census. Future work with the sources on the local level, comparing several censuses and other source material for the same localities and persons will be necessary in order to get a fuller picture of the degree to which ethnicity was marked erroneously.

## 6. Constructed Ethnicity Variables

We shall now outline how to construct new ethnicity variables in order to harmonize the differences over time and space investigated above. First let us summarize how censuses during the period varied along several dimensions:

- 1) Collection of Data on Head of Household vs. on Individuals
- 2) The Respondent's Native Language
- 3) Data Collection by Census Takers vs. Self-Administration
- 4) Under-Enumeration
- 5) The Wording of the Variable: Nationality – Ethnicity – Race
- 6) Individual Ancestry vs. Classification on Language
- 7) Reporting and Documentation Practice
- 8) Number of Ethnic Categories
- 9) Changing Borders.

With the techniques used in the North Atlantic Population Project and other census projects (cf [ipums.org](http://ipums.org)), the censuses can still be made comparable over time and space. A most important task is to construct group level variables for each individual giving the predominant ethnic group in each locality. It is necessary to construct several such variables denoting dominant ethnicity both on the municipality, ward and farm or place level. These will be made available and documented on the web at [nappdata.org](http://nappdata.org). A difficult question is whether to assign ethnicity on the basis of older source material in cases where newer sources lack ethnicity markers for the whole or parts of the population. While in many places ethnicity seems quite stable over the decades or even centuries, the situation can be very dynamic as is seen along the US western frontier or when Norwegians took over rich fishing places in Sami fjords. Thus, such “ahistoric” construction must be performed by researchers with the necessary local history insight. It should be less controversial to copy the ethnic markers from the contemporary ethnographic maps made by Friis as a constructed variable on the farm or settlement level in the 1865 census (Cf [www.dokpro.uio.no/friis-kartene.html](http://www.dokpro.uio.no/friis-kartene.html)).

Another dynamic aspect of ethnicity that should be handled more adequately by the constructed variables is the mixing of ethnic groups due to intermarriage and migration. Much information is lost when large groups of individuals are simply called “mixed” whatever constituent ethnic groups are involved. We must check our data thoroughly when we find new ethnic elements in what used to be a homogeneous ethnic locality. Especially useful would be one or more constructed ratio scale variables giving the proportion of people on each farm belonging to the dominant ethnic group (s) in the locality. A variable giving the proportion of individuals with mixed ethnicity will also be useful since it can be used to weight the ethnicity factor in statistical analysis. A further question is to what extent any constructed variable should distinguish between cultural and ancestral factors. At least a partial answer to this dilemma might be to construct group level language variables in cases where such information is available. It goes without saying that the constructed variables about ethnicity need to be documented thoroughly for each nation and even province, and with abundant references also to local and regional literature.

## 7. Conclusion

From 1845 to 1930 the decadal censuses for Northern Norway contained the most detailed information on ethnicity among all the censuses in Western Europe. While ancestry was the basis for the allocation of the ethnic markers in the first of these censuses, more weight was put on cultural criteria such as language towards the end of the nineteenth century. The nominative 1865, 1875 and 1900 census manuscripts have been transcribed and encoded so that they are available for analysis on the individual level in a system containing also several other censuses from the nations around the North Atlantic, and the 1910 census data will soon follow suit. Tests reveal that although the definition of ethnicity changed over time, the census takers were aware of the problems and in general managed to allocate the ethnic markers, including the distinction between the Finnish and the Sami who were often called Fins. The article concludes by outlining a number of constructed variables which may enhance the researchers' attempts to compare ethnicity over time and across space.

## References

- Aubert, Vilhelm (1978): Den samiske befolkning i Nord-Norge. Oslo: Statistisk sentralbyrå. ([http://www.ssb.no/histstat/art\\_107.pdf](http://www.ssb.no/histstat/art_107.pdf)).
- Barth, Fredrik (1969): Ethnic groups and boundaries: the social organization of culture difference. Bergen: Universitetsforlaget.
- Bjørklund, Ivar (1985): Fjordfolket i Kvænangen: Fra samisk samfunn til norsk utkant 1550-1980. Tromsø: Universitetsforlaget.
- Drake, Michael (1991): Ulikhetens mønster: Tromsø by i 1860-åra. Tromsø: University of Tromsø, Institute of History, 25.
- Eriksen, Knut Einar / Niemi, Einar (1981): Den finske fare: sikkerhetsproblemer og minoritetspolitikk i nord 1860-1940. Oslo: Universitetsforlaget.
- Handlin, Oscar (1951): The uprooted: the epic story of the great migrations that made the American people. Boston, Mass.: Little, Brown and Company.
- Hansen, Lars-Ivar (1994): Samene i forrige århundres folketellinger. Registreringspraksis i Astafjord prestegjeld 1865-1900. In: Festskrift til Ørnulf Vorren. Tromsø: Tromsø museum, University of Tromsø.
- Hansen, Lars-Ivar / Meyer, Tore (1991): The ethnic classification in the late 19th century censuses. A case study from Southern Troms, Norway. In: Acta Borealia (2).
- Hauglid, Anders Ole (1981): Balsfjorden og Malangens historie. Balsfjord kommune.
- Hobsbawm, Eric (1990): Nations and nationalism since 1780: programme, myth, reality. Cambridge: Cambridge University Press.
- Isaksen, Kristin (1988): Folketellinger 1815-1980. In: Arkivmagasinet (2). 21-29.
- Jåstad, Hilde Leikny (2003): Spedbarnsdødeligheten i Tana i perioden 1840-1914. Tromsø: University of Tromsø, Institute of History.
- Kiær, Nicolai (1882): Bidrag til en norsk Befolkningsstatistik, Oslo: Statistics Norway/Statistisk sentralbyrå.

- Lie, Einar and Sissel Boquist. (2001): Faktisk talt: statistikkens historie i Norge. Oslo: Universitetsforlaget.
- Manning, Peter (2005): Migration in World History. New York: Routledge.
- McCaa, Robert / Ruggles, Steven (2001): The Census in global perspective and the coming microdata revolution. In: The 14th Nordic Demographic Symposium. Tjøme: Nordic Demographic Society.
- Minde, Henry (1975): Folketilveksten i Alta 1690-1825: Belegg – utvikling – bakgrunn. Tromsø: University of Tromsø, Institute of History.
- Myhre, Jan Eivind (2003): I nasjonalstatens tid, 1814-1940. Oslo. Pax forlag.
- Niemi, Einar (1977): Oppbrudd og tilpassing, Vadsø kommune.
- Niemi, Einar (2000): Når man skyter hunder. En debatt om reindrift i Finnmark ved midten av 1800-tallet. In: Den komplekse kontinenten. Staterna på Nordkalotten och samerna i et historiskt perspektiv. Sköld, P / Lantto, P (eds.). Umeå.
- Ones, Per. (1970): Den samiske nasjon. Oslo: Pax. Electronic version at <[http://www.nb.no/utlevering/contentview.jsf?urn=URN:NBN:nonb\\_digibok\\_2007041102001](http://www.nb.no/utlevering/contentview.jsf?urn=URN:NBN:nonb_digibok_2007041102001)>.
- Roberts, Evan (2003): The North Atlantic Population Project: An Overview. In: Historical Methods 36 (2). 80-88.
- Patriarca, Silvana (1994): Statistical Nation Building and the Consolidation of Regions in Italy. In: Social Science History 18 (3). 359-376.
- Schøyen, Carl (1918): Tre stammers møte. Kristiania: Gyldendal.
- Sogner, Sølvi (1979): Folkevekst og flytting. Oslo: Universitetsforlaget.
- Statistics Norway (1868-69): Resultaterne af Folketællingen i Norge i Januar 1866. Kristiania.
- Statistics Norway (1878-81): Resultaterne af Folketællingen i Norge i Januar 1876. Kristiania.
- Statistics Norway (1894-98): Folketællingen i Kongeriget Norge 1891. Recensement du 1er Janvier 1891. Kristiania: Det statistiske centralbureau.
- Statistics Norway (1906): Folketællingen i Kongeriget Norge 3 December 1900: Hovedoversigt. Recensement du 3 Décembre 1900: Aperçu General. Kristiania, Statistisk centralbyrå.
- Statistics Norway (1916): Folketællingen i Norge 1 december 1910: Hovedoversigt. Kristiania: H. Aschehoug & Co.
- Statistics Norway (1995): Historisk statistikk 1994. Historical statistics 1994. Oslo: Statistisk sentralbyrå.
- Statistisches Institut, Internationalen (1887): 1. Thema: Die Verhandlungen und Beschlüsse des Internationalen Statistischen Instituts in Betreff einer einheitlichen Aufarbeitung der Volkszählungen. Arbeiten der Demographischen Section, Wien: Internationalen Statistischen Institut.
- Steinlien, Øystein (1986): Kulturell endring og etnisk kontinuitet. In: Kulturer i norr. Svanberg, I. / Lyngby L. (eds.) NEFA-Norden. 49-54.
- Thorsen, Hanne Christine (1972): Registreringen av den samiske befolkning i Nord-Norge fra 1845 til 1970. Mastergradsavhandling i sosiologi. University of Oslo.
- Thorsen, D. (1865): Om Forholdene i Tanens Lægedistrikt, af Distriktslæge Thorsens medicinalberetning for 1865. NOS C No 4. Kristiania.

- Thorvaldsen, Gunnar (1995): Migrasjon i Troms i annen halvdel av 1800-tallet. En kvantitativ analyse av folketellingene 1865, 1875 og 1900. Tromsø: University of Tromsø.
- Thorvaldsen, Gunnar (1997): On Boundaries and Areas in Local Historical Research. In: Festskrift til Sivert Langholm. Myhre, J. / Kjeldstadli, K. (eds). Oslo: The Norwegian Historical Association. Electronic version in English at <http://www.rhd.uit.no/art/boundaries.html>.
- Thorvaldsen, Gunnar (2004): Koloniseringen av Bardu og Målselv. In: Folketellinger gjennom 200 år. Soltvedt, K. (ed). Oslo: Statistisk sentralbyrå. 48-70.
- Thorvaldsen, Gunnar (2006): Away on Census Day. Enumerating the Temporarily Present or Absent. In: Historical Methods 39 (2). 82-96.
- Thorvaldsen, Gunnar (2007): An International Perspective on Scandinavia's Historical Censuses. In: Scandinavian Journal of History 32 (3). 237-257.
- Thuen, Trond (1987): One community – one people? Ethnicity and demography in a North-Norwegian community 1865-1930. In: Acta Borealia 23 (1). 65-83.
- Torp, Eivind (1986): Registrering av etnisitet i folketellinger. In: Heimen 23 (2). 67-77.
- Torp, Eivind (1990): Information about Information: Relating different historical sources to one another. In: Acta Borealia 26 (2). 81-85.
- Turner, Fredrik Jackson (1893/1966): The significance of the frontier in American history. Ann Arbor. Reprint 1966 of annual report of the American historical association for the year 1893.
- Westergaard, Harald. (1932): Contributions to the history of statistics. London.